

SEQUENCE LISTING - WEE virus strain 71V-1658 - SEQ ID NO 1

LOCUS AF214040 11484 bp RNA VRL 27-JAN-2000
 DEFINITION Western equine encephalomyelitis virus strain 71V-1658
 nonstructural polyprotein and structural polyprotein genes,
 complete cds.
 ACCESSION AF214040 AF143811
 VERSION AF214040.1 GI:6760410
 KEYWORDS .
 SOURCE Western equine encephalomyelitis virus.
 ORGANISM Western equine encephalomyelitis virus Viruses; ssRNA
 positive-strand viruses, no DNA stage; Togaviridae; Alphavirus.
 REFERENCE 1 (bases 1 to 11484)
 AUTHORS Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A.,
 Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.
 TITLE Complete genomic RNA sequence of western equine
 encephalitis virus and expression of the structural genes
 JOURNAL J. Gen. Virol. 81 Pt 1, 151-159 (2000)
 MEDLINE 20109043
 REFERENCE 2 (bases 1 to 11484)
 AUTHORS Netolitzky,D.J., Schmaltz,F.L., Parker,M.D., Rayner,G.A.,
 Fisher,G.R., Trent,D.W., Bader,D.E. and Nagata,L.P.
 TITLE Direct Submission
 JOURNAL Submitted (08-DEC-1999) Medical Countermeasures Section,
 Defence Research Establishment Suffield, P.O. Box 4000, Stn Main,
 Medicine Hat, Alberta T1A 8K6, Canada
 FEATURES
 source 1..11484
 /organism="Western equine encephalomyelitis virus"
 /strain="71V-1658"
 /db_xref="taxon:11039"
 5'UTR <1..24
 misc_feature 1..11484
 /note="sequence of this entry added to the sequence of
 AF214039, strain CBA87, constitutes the Western equine
 encephalomyelitis virus complete genome"

 CDS 25..7428
 /note="encodes nsP1, nsP2, nsP3 and nsP4 proteins"
 /codon_start=1
 /transl_except=(pos:5581..5583,aa:OTHER)
 /product="nonstructural polyprotein"
 /protein_id="AAF28339.1"
 /db_xref="GI:6760411"

 /translation="MERIHVDLDADSPYVKSRLQRTFFQFEIARQVTDNDHANARAFS
 HVATKLISEVDRDQVILDIGSAPVRHAHSNHRHYHCICPMISAEDPDLRQRYAERLKK

FEB 27 1968

NEQICLYILGESMSSIRSKCPVEESEASAPPHTLPCLCNYAMTAERVYRLRSKKEQF
 AVCSSFLLPKYRITGVQKLQCSKPVLFSGVVPVAVHPRKYAEIILETPPPPTTTVIC
 EPTVPERIPSPVISRAPSAESLLSLGGVSFSSSATRSSTAWSDDYDRRFVVTADVHQAN
 TSTWSIPAPGLDVLQPSDVTDSHWSIPSASGFEVVRTPSVQDLTAECAKPRGLAIBMQ
 DFNTAPFQLSDYRPVPAPRRRPIPSPRSTASAPPVVKPRRTKYQQPPGVARAISEAE
 LDEYIRQHSNXRYEAGAYIPSSSETGQGHLLQKQSVRQCKLQEPILDRAVHEKYYAPRLD
 LEREKMLQKKLQLCASEGNRSRYQSRKVENMKAITAERLISGLGTYLSSEVNPVECYR
 VNYPVPIYSSTVINRFTSAEVAVKTCNLVIQENYPTVASYICITDEYDAYLDMVDGASC
 CLDTATFCPAKLRSYPKKHSYLQPEIRSAVPSPIQNTLQNVLAATKRCNCVNTQMREL
 PVLDSAAFNVDCFKKYACNDEYWDTFRDNPIRLTTENTVQYVTCLKGPKAAALFANTH
 NLKPLQEIIPMDQFVMDLKRDKVKTPTGKHTTEERPKVQVIQAADPLATAYLCGIHREL
 RRLNAVLLPNIHITLFDMSAEDFDAIIAEHFHHGDPVLETDIASFDKSEDDAIAISALM
 ILEDLGVDQPLLDLIEAAGNITSVHLPTGTRFKFGAMMKSGMFLTLFVNTLVNIMIA
 SRVLRERLTTSAACAISIGDDNIVHGVVSDTLMAERCATWLNMEVKIIDAVIGIKAPYF
 CGGFILVDQITGTACRVADPLKRLFKLGKPLPVDDTQDCDRRRALHDEAMRWNIRGIT
 DELVKAVESRYEIIILAGLIITSLSTLAESVKNFKSIRGSPITLYG"

CDS 7473..11183
 /note="encodes nucleocapsid, E3, E2, 6K and E1 proteins"
 /codon_start=1
 /product="structural polyprotein"
 /protein_id="AAF28340.1"
 /db_xref="GI:6760412"
 /translation="MFFYPQLNFPPVYPTNPMAYRDPNPPRRRWRPFRPLAAQIEDL
 RRSIVNLTFRQSRPNPPPGPPPKKKKSAPKPKPTQPKKKKQQAQKRTKRKPKPGKRQRM
 CMKLESDKTFFPIMLNGQVNGYACVVGGRMLKPLHVEGKIDNEQLAAVKLKKASMYDLE
 YGDVPQNMKSDTLQYTSDKPPGFYNWHHGAVQYENGRFTVPRGVGGKGDSCRPIILDNR
 GRVVAIVLGGANEGRTALSVVTWNQKGVTIIRDTPEGSEPSWLSLTALCVLSNVTFFPCD

KPPVCYSLTPERTLDVLEENVNDPNPYDTLLENVLKCPSSRRPKRSITDDFTLTSPYLG
 CPYCRHSTPCFSPIKIENVWDESDDGSIRIQVSAQFGYNQAGTADVTKFRYMSFDHHD
 DIKEDSMEKIAISTSGPCRRLLGHKGYFLLAQCPPGDSVTVSITSGASENSCTVEKKIR
 RKFVGREEYLFPPVHGKLVKCHVDHLKETSAGYITMHRPGPHAYKSYLEEASGEVYI
 KPPSGKNVTYECKCGDYSTGIVSTRTKMNGCTKAKQCIAYKSDQTKWVFNSPDLIRHT
 DHSVQGLKHIPFRLTPTVCPVPLAHTPTVTWKFGITLHLTAMRPTLLTTRKLGLRAD
 ATAEWITGSTRNFSVSGREGLEYVWGNHEPVRVWAQESAPGDPHGWPHEIIHHYYHRH
 PVTYIVILCGLVALAILVGTASSAACIAKARRDCLTPYALAPNATVPTALAVLCCIRPT
 NAETFGETLNHLWFNNQFLLWAQLCIPLAALVILFRFCFSCCMPFLLVAGVCLGKVDAP
 EHATTVPNVPGIPIYKALVERAGYAPLNLEITVVSSELTPTSTNKEYVTCKFHTVIPSQ
 VKCCGSLECKASSKADYTCRVFGGVYPFMWGGAQCFCDSSENTQLSEAYVEFAPDCTID
 HAVALKVHTAALKVGLRIVYGNTTAHLDTFVNGVTPGSSRLDKVIAGPISAAFPSFDDH
 KVVIRKGLVNYDFPEYGAMKPGAFGDIQASSLDATDIVARTDIRLLKPSVKNIHVPI
 TQAVSGYEMWKNNNGRPLQETAPFGCKIEVEPLRASNCAYGHIPISIDIPDAAFVRSS
 ESPTILEVSTVADCIYSADFGLSLTLQYKADREGHCFVHSHSTTAVLKEATHTVAV
 GSITLHFSTSSPQANFIVSLCGKSTCNAECKPPADHIIIGEPHKVDQEFQAQAVSKTSW
 NWLLALFEGGASSLIVGLIVLVCSSMLINTRR"

3'UTR

11184..11484

BASE COUNT 3251 a 2842 c 2808 g 2583 t
 ORIGIN

1	accctacaaa	ctaategac	caatatggaa	agaattcacg	ttgacttaga	tgctgacagc
61	ccgtatgtca	agtcgtttaca	gcggacgttt	ccacaatttg	agatcgaagc	aaggcagggtc
121	actgacaatg	accatgccaa	tgccagagcg	ttttcgcatg	tggaacacaa	gctcattgag
181	agcgaagtcg	accgggacca	agttatcttg	gacattggaa	gtgcgcccgt	cagacatgca
241	cattccaate	accgctatca	ttgtatctgc	cctatgataa	gcgctgaaga	ccgggacaga
301	ctacaacggt	atgcagaaaag	acttaagaaa	agtgacatta	ccgacaagaa	catagcctct
361	aaggcggcag	acctgctgga	agtcattgtca	acaccagacg	cagagactcc	atctctgtgt
421	atgcacacag	acgccacgtg	taggtacttt	ggaagtgtag	cagtatacca	agatgtgtac
481	gcagtcacatg	caccgacatc	aatctaccac	caggcgctta	aaggagttag	gacaatttac
541	tggataggct	ttgacacgac	cccttttatg	tacaaaaaca	tggcagggttc	ctaccctact

601	tacaacacga	actgggctga	cgagagagta	ttggaagcac	gtaacattgg	cctcggtaac
661	tcagatcttc	aggagagcag	gcttggaaaa	ctctcaatcc	ttaggaagaa	gaggtccaa
721	cctactaata	agatcatatt	ctcggttggt	tcaacatct	acacagaaga	tagatcactg
781	ttacgtagct	ggcatcttcc	aaacgtgttc	cacttgaaag	gaaagtctaa	cttcacaggt
841	agatgtggga	ccattgtcag	ctgtgaaggg	tacgtcatca	aaaagataac	gatcagccca
901	ggactatacg	gtaaaagtga	gaacttggcg	tcacacaatg	atcgcgaggg	ttctcttgat
961	tgcaaaagct	catagacgtc	gcgcggcgag	agggtttctt	ttgctgtgtg	tacgtatgta
1021	ccagccacac	tttgcgcatc	gatgacaggg	attctggcaa	ctgacgttag	tgtgggatgac
1081	gcacaaaaac	tatttggttgg	gctcaaccaa	aggattgtcg	tcaatggtag	gacgcaaaga
1141	aatactaaca	caatgcagaa	ctatctatta	ccagtggtcg	cccaggcggt	ttccaggtgg
1201	gcgcgtgaac	atcgtgccga	cttggacgac	gagaaagaac	taggggtgcg	ggagcgcact
1261	cttactatgg	gctgtgctgt	ggctttcaag	acccagaaaa	tcacatccat	ctacaagaag
1321	cctggtagcg	aaacaattaa	gaaagtagct	gccgtctttg	actcattttg	gattccacgc
1381	cttaccagcc	acgggctcga	tatgggcttc	cgccgtaggg	tcaagctgct	gcttgaacca
1441	actgtcaaac	cgcgaccggc	tattacaatg	gccgatgtgc	agcatctggc	tggcttaca
1501	caagaagctg	aagaagtgcc	tcagcgggaa	gagatcacag	aagccctgcc	acccttgctc
1561	cctgaaatag	aaaaagagac	cgtagaggca	gaagttagacc	tcattatgca	agagcgagga
1621	gcaggtagcg	tggagacacc	acgaggacac	atcaggggtga	caagttaccc	agggcgaaag
1681	aagatttgggt	cttaccgtat	actttcaccc	caggcgggat	tgaatagtga	aaaactggga
1741	tgtatccacc	cttggcgcca	acaagtagct	gtaattgactc	acaaaggtag	ggcggtgcga
1801	tacaaagtgc	agccatacca	cggtaaagtc	attgtaccag	aagggacggc	ggccctgttt
1861	caagacttcc	agggacttgag	tgagagcgct	acgatcggtt	tcaacgagag	ggagttcgta
1921	aacagatacc	tggaccatcc	cgcaatcaac	ggaggagcgc	taaacactga	cgaagaagtac
1981	tataagactg	taaagactca	ggacacagac	tcagaatacg	tcttcgatat	tgacgcacga
2041	aaagtgtgta	agcgacaaga	cgcaggtccc	ttgtgcctac	ccgggtgactc	ggtagatcca
2101	ccatttccag	agtttgcgta	cgagagttct	aagacacgac	cagcagaccc	tcacaaagtc
2161	ccaaccatcg	gagttctatg	agtgccaggt	tcaggtaaat	ctggaatcat	caaaagcgct
2221	gtgactaaga	aagatctggg	tgtgagtgcg	aagaagggaaa	actgcgcaga	aatcatcagg
2281	gatgtaagga	ggatgagacg	tatggatggt	gctgctagga	ctgtcgattc	agtgcttcta
2341	aatgggggta	agccaccctg	taacactctg	tacatttagt	agggcattgc	ctgccatgca
2401	gggacgctgc	tggcactgat	tgccatcgct	aaacctaaaga	aagtggttat	gtgcggggag
2461	ccaaaacaat	gcggcttctt	taacatgatg	tgccgtgaaag	tacattttta	ccatgacata
2521	tgcactgaag	tgtaccataa	aagcatctct	aggaggttca	cacagactgt	aaccgcctac
2581	gtctccacgc	tcttctacga	caagcgaatg	aagacgggta	acccattgtc	tgataaaaatc
2641	atcatagata	ccacagggag	cacaaaagcg	cacaaagatg	atctgattct	aacctgtttc
2701	agagagtggtg	tgaaacagct	acagattgac	tacaaaatcc	acgaatctac	gactgcggct
2761	gcactgcgaag	cgactacggc	gaaagcgctg	tatgctgtca	ggtaacaaag	caacgagaat
2821	ccactctact	gcagactctc	tgagcactgt	aacgtgtgta	ttacacgcac	agaaaaacgc
2881	attgtctgga	agacgctagc	tggatgaccc	tggataaaga	cacttacagc	taaatatccc
2941	ggggatttca	cggtcttact	ggacgactgg	cagcgccga	acgacgccat	tatggcacgc
3001	gttcttgata	agccgcgacg	agctgatgtg	ttccagaata	aggtgaaagt	cttgcggggc
3061	aagggctttag	agccagctct	ggccacggcc	aacatttgtc	tgacgagaca	gcagtgggag
3121	acgttgaccc	catccaagca	tgacagagcg	tactcacctg	aaatggcact	gacactcttt
3181	tgcaccaggt	tctttggagt	agacctggac	agtggggtat	ttccgcgtcc	taccgtcgca
3241	cttactttaca	gggatcagca	ctgggataac	tcggcagggga	agaacatgta	tgggcttaat
3301	agagaggtga	caaaggagtt	gtcacggcga	tatccgtgca	tcacaaaagc	ggttgacaca
3361	ggcagggtag	ctgatataag	gaataatacc	atcaaggctt	actctccaac	aataatgtg
3421	gttccattaa	ctgcgcgggt	gccccactgc	ttgatcggtg	accacaaagg	acagggtaaca
3481	actgatcaca	gcggattcct	atctaagatg	aagggcaaat	ctgtgttggt	gatcggcgat
3541	cctatcacga	ttccagggaa	gaaagtagag	tccatgggtc	cattgcccac	taataccatc

3601	agggtgtgac	tcgatttggg	aatacctage	catgtcggta	aatatgacat	tatctttgtc
3661	aatggttagga	cccogtagac	gaaccatcac	taccaacagt	gcgaggatca	cgctatccac
3721	cacagcatgc	taacgtgtaa	ggctgtccac	cacctgaaca	ctggcggaac	atggtgtggct
3781	atagggtatg	ggctgtgtga	tcgcgcgaac	gagaatatca	tcaactcggt	ggcagcgctca
3841	tttaggttcta	ccogtgtctg	tcagcctaag	aacactgcgc	aaaactatga	ggttctcttc
3901	gtgttctctgc	gcaaggacaa	cggcaaccac	acacatgacc	aggacagact	cggtgtagtg
3961	cttgacaaca	tctatcaagg	gtcaaccagg	tacgaggcag	ggagagctcc	agcgtagaca
4021	gtgatcagag	gtgacattag	caagagcgct	gaccaacgta	tcgttaatgc	gcttaaatagc
4081	aaagggtcaac	caggtttccgg	agtgtgcggt	gcactgtacc	gaaaaatggc	ggctgctttt
4141	gatagacagc	caatagctgt	cgggacggct	agacttgtga	agcaccgaac	gctcatcata
4201	catgctgtag	gacccaattt	ttctaagatg	ccggaaccgg	agggcgacct	taagctcgca
4261	gctgcctaca	tgagcatagc	gtccatcgct	aacgctgagc	ggattacaaa	aatatcagta
4321	cgctactgtg	caaccggcat	ctattctggg	ggcaaaagtc	gagtgatgca	atcattgcat
4381	cacctgttca	ctgctttcga	cactacggat	cgcgatgtca	ccatatattg	ctgggataaa
4441	caatgggaga	ccagggaatc	cgaggccatt	cccgcaaaag	aaagcgctga	aattctggat
4501	gatgacaagc	cagtagacat	tgacttggct	agggtccacc	caaacagctc	tttggcaggc
4561	agaccagggt	actccgtcaa	tgagggcaag	ttgtatctat	acctgggaag	tacacgacttc
4621	catcagaccg	ccaaggacat	tgccgaatc	catgcaatgt	ggcccaacaa	attctgaggt
4681	aatgagcaga	tttgcttgta	catcctgggg	gagagtatgt	ccagatcccg	ctccaaatgc
4741	ccagtagagg	agtccagagg	gtctgtctca	ctccacacac	ttccatagct	gtgtaattac
4801	gctatgacgg	ctgagccggt	atcacggttg	cgctctgcga	agaaaagcct	gttgcgcgta
4861	tgctcatcat	tctctgttgc	gaagtacagg	atcacaggcg	tgcagaagct	acagtgcaagc
4921	aaaccagctcc	tgttttcagg	cgctgtacca	cgcgctgtac	accocaggaa	gtacgctgaa
4981	ataattctag	aaacgcaccc	accgcagcag	acgacaaccg	taatatgtga	accaactgtg
5041	ccagaacgta	taccagctcc	gggtgatttc	agagcaccaa	gtgcggaaatc	actgctatcg
5101	cttggcgccg	tctcgttctc	tagctctgcc	acacgctcgt	caaccgcctg	gagcgactat
5161	gacaggcggt	ttgtggttac	agctgatgtg	catcaagcga	acacgtctac	gtggagcatc
5221	cctagtgtctc	ctggccttga	cgctccagctg	cctctgacg	tcactgattc	ccactggagt
5281	attccaagtg	catcaggctt	tgaagtgaga	acaccatctg	tacaggacct	aactcgcgag
5341	tgtgcgaagc	ctcgtggact	ggccgaaata	atgcaagact	tcaatatctg	tcctttccag
5401	tttctttcgg	actacagacc	agtagccgga	ccacgggagc	gccccatccc	atcacctaga
5461	tcgacggctt	ccgcacctcc	agttccaaag	ccacgcagga	ctaagtacca	acaaccacca
5521	ggagtgcgta	gagcgtatc	agaagcggag	ttggacgagt	acatccgtca	acactccaac
5581	tgacgggtat	aagcgggagc	gtatattttc	tcatcggaag	caggcccaag	tcaccttcaa
5641	cagaaatcag	tacgtcaatg	taaaactaaa	gaacctatat	tggatcgggc	cgctccatgag
5701	aggtattacg	ccocgcgcct	cgatctcgaa	agagagaaaa	tgttacagaa	atcacctcaa
5761	ttatgcgcct	ctgaaggaaa	tagaagcagg	tatcaatcac	gaaaagtaga	aaatatgaaa
5821	gcaattacag	cggagcgact	caatttctga	ttgggacatc	atctatcatic	agaagtgaat
5881	cctgtcgagt	gttacagagt	caattatcct	gtaccaatct	actcgtaaac	ggtaattaac
5941	aggtttacat	ctgcagaggt	cgcggttaaa	acgtgcaact	tagttatcca	agagaattac
6001	ccacagtagc	ccagttattg	tataacagat	gaatacagtg	cgtaatttga	ctggtggagc
6061	ggcgctatcg	gtgctctaga	tacagccact	ttttgtccgg	tcaaacatgag	aagctacca
6121	aaagaacata	cgctatttgc	gccagagata	agatcagcgc	ctccactcgc	tatacagata
6181	acattacaaa	atgtatttgc	tcagcgtact	aaaaggaatt	gcaacgttac	ccaaatgcga
6241	gaattacctg	tcttagattc	ggcggtcatt	aatgttgatt	gtttcaagaa	ataccgcatgc
6301	aatgatgagt	actgggatac	ctttcgcgat	aacctatcca	ggtaaacctac	agagaacggt
6361	acgcaatatt	tgacaagaat	gaaagggccg	aaagcagcag	catgttttgc	gaatactcat
6421	aatctaaaac	cgcttcaggga	gataccaatg	gatcaattcg	tcatggatct	aaagagagat
6481	gtcaaaagtta	ctcccggcac	gaaacatata	gaggagcggc	ctaaggtgca	ggttattcag
6541	gctgcagatc	cccttgctac	cgcttacctt	tgccggatcc	atcgggaaat	agtcogtaga

6601	ctgaatgcggg	tgcctctcgcc	aaatatccat	actctctctcg	acatgtcagc	ggaagatttt
6661	gatgcgatta	ttgctgtaaca	tttccaccac	ggcgaccocag	tatttgaaac	ggacatcgcg
6721	tcggtttgata	aaagcgaaga	cgacgcgtatc	gccatttctcg	cgttgatgat	ccttgaggac
6781	ttaggtgtcgc	accaaccgcgt	cttagatttg	atagaggcg	cgttcgggcaa	tatcacatct
6841	gtgcacctac	ctacaggaac	gaggttttaaa	ttgttgcca	tgatgaaatc	cggtatgttc
6901	ttaacgctgt	ttgtcaacac	actagtcfaat	atcatgatgt	ctagcagagt	actacgtgaa
6961	cgggttaacca	cgtcagcgtg	cgcgccctct	atcggcgcag	ataacatagt	gcatggtgtg
7021	gtctccgaca	ccttgatggc	ggagagatgc	gccacttgcc	tgaacatgga	agtaaaaaatt
7081	attgatgcag	ttattggtat	caaagcaccc	tacttctgtg	ggggatttat	cctggtggac
7141	cagataacag	gcacagcctg	cagagtgcga	gacctctaa	aaaggctttt	taagcttgga
7201	aaaccattgc	cagtcgatga	taccacaagac	tgcgaccgcc	gccgggcact	gcatgatgaa
7261	gcaatgcatg	ggaacagatg	tggaattacg	gacgagttag	tgaaggccgt	agaatccaga
7321	tacgagatca	tactggcagg	cctgatcctc	acgtctctgt	ccacgttagc	cgaaagcgtt
7381	aagaacttca	agagcataag	agggagccca	atccacctct	acggctgacc	taaatagggt
7441	acgtagtaga	cacgcacctc	cccaccggca	gaatgtttcc	ataccctgac	ctgaaacttc
7501	caccagttta	ccctacaaat	ccgatggctt	accgagatcc	aaacctctct	aggcgccgct
7561	ggaggccggt	tcgggccccg	ctgggtgctc	aaatcgaaga	tcttaggagg	tcatagtgca
7621	acttgacttt	caaacacaga	tcacctaatc	cgccgcccagg	tccaccgcga	aagaagaaga
7681	agagtgctcc	taagcctaaa	cctactcagc	ctaaaaagaa	gaagcagcaa	gccaaaggga
7741	cgaacgcgca	gcctaaacca	gggaaacgac	accgtatgtg	tgaacatagt	gactcggaca
7801	agacatttcc	gatcatgtcg	aacggcccaag	tgaatggata	tgccctgcgt	gtcggaggaa
7861	ggctgatgaa	accaactcac	gttgaggaa	aaattgataa	tgagcaattt	gcccgcgtga
7921	aattgaagaa	ggctagactg	tacgacttgg	agtcaggcga	cgttcccgac	aacatgaaat
7981	cagacacgct	gcagtacacc	agcgacaaac	caccgggctt	ctacaactgg	caccacggcg
8041	cagtcagcta	tgaagattgg	agatttaccg	taccgagagg	agtggggcgg	aaaggcgaca
8101	gcggaagacc	gatccctggac	aacagaggga	gagttgtggc	tattgttcta	ggagggtgcaa
8161	atgagggcac	gcgtacggcg	ctttcagtg	tcacttggaa	ccagaaagg	gtgaccattt
8221	gggatacccc	cgaaggttct	gaaccgttgt	cactagttac	agcgcctatg	gtgcttccga
8281	atgtcacggt	cccatgcgac	aaaccaccgg	tgtgctattc	actgacgcga	gaacgaacac
8341	tcgacgtgct	cgaagagac	gtcgacaatc	caaattacga	cacgctgtgt	gagaaactct
8401	tgaattgtcc	atcacgcggg	cccaaacgaa	gcattacaga	tgacttccga	ctgaccagtc
8461	cctacactgg	gtcttgcccc	tattgcagac	actcaacgcg	gtgtttcagc	ccaataaaaa
8521	ttgagaaact	gttggcagaa	tctgatgatg	gatcgactag	aatctcagtc	tcggcacaat
8581	tcggctacaa	tcaggcaggc	actgcggatg	tcaccaaatt	ccgtttacatg	tctttcgacc
8641	acgacatgaa	catcaaggaa	gacagtatgg	agaaaaatgc	tatcagcaca	tctggaccct
8701	gcgctgcgtc	tgggccacaaa	gggtacttcc	tgttagctca	atgtctccca	ggtgacagt
8761	taaccgctcag	tatcacgcagc	ggagcatctg	agaattcatg	caccggtggag	aaaaagatca
8821	ggagggaagt	tgtcggtaga	gaggagtact	tgttcccgac	cgctcccgga	aagctggtaa
8881	agtgcacagt	ttacgatcac	ttgaaggaga	cgtctgcggg	gtacataacc	atgcacaggc
8941	caggcccaaca	cgcgtataag	tcctatctgg	aggaagcgct	aggcgaagt	tacattaaac
9001	cacctctcgg	caagaacgtc	acctacgaat	gtaagtgtgc	cgactacagc	gctgacagt
9061	tgagcagcgc	cagcaagatg	aacggctgca	ctaaagcaaa	acagttgcatt	gcctacaga
9121	gcgaccaaac	gaattgggtc	ttcaactgcg	cggatcttat	taggcacaca	gaccactcag
9181	tgcagggtaa	attgcacatt	ccattccgct	tgacaccgac	agtctgcccg	gttcggttag
9241	ctcacacgcc	tacagtcacg	aagtgggtta	aaggcatcac	cctccacctg	actgcaatgc
9301	gaccaacatt	gctgacaacg	agaaaaattg	ggctgcgcagc	agacgcaaca	cagaatgga
9361	ttacagggtg	tacatccagg	aatttttctg	tggggcggaga	aggcgtggag	tacgtattgg
9421	gtaacactga	accagtacga	gtctggggcc	aggagtgcgc	accggcgac	cpacatggat
9481	ggcgcgatga	gatcatcatc	cactattatc	atcgccatcc	agtctacact	gtcatttgtc
9541	tgtgtggtgt	cgtctctgct	atccttggtag	gcactgcact	atcagcagct	tgcattcgca

SEQUENCE LISTING - pcDWXH-7 - SEQ ID NO 2

<210> SEQ ID NO 2
 <211> LENGTH: 4150
 <212> TYPE: DNA
 <213> ORGANISM: WESTERN EQUINE ENCEPHALITIS VIRUS
 <214> FEATURES:

Location/Qualifiers

source 1..4150
 /organism="Western equine encephalomyelitis virus"
 /strain="71V-1658"
 vector sequence 1-9
 5' SacI primer 9-20
 3' end of NS4 gene 16-114
 intragenic region 115-159
 CDS 160-3869
 /note="encodes capsid, E3, E2, 6K and E1 proteins"
 /codon_start=160
 /product="structural polyprotein"

/translation="MFPYPQLNFFPVYPTNPMAYRDPNPPRRRWRPFRPPLAAQIEDL
 RRSIVNLTfKQRSPNPPPGPPPKKKKSAPKPKPTQPKKKKKQAKRTKRKPKPGKRQRM
 CMKLESDKTFPIMLNGQVNGYACVVGGRMLMKPLHVEGKIDNEQLAAVKLKKASMYDLE
 YGDVPQNMKSDTLQYTSDKPPGFYNWHGAVQYENGRFTVPRGVGGKGDSCRPIILDNR
 GRVVAIVLGGANEGTRTALS SVVTWNQKGV TIRDTPEGSEPSLVTALCVLSNVTFPCD
 KPPVCYSLTPERTILDVLEENVDPNPNYDTLLENVLKCPSSRRPKRSITDDFTLTSPYLGF
 CPYCRHSTPCFSPIKIENWDESDDGSIRIQVSAQFGYNQAGTADVTKFRYMSFDHHD
 DIKEDSMEKIAISTSGPCRRLGHKGYFLLAQCPPGDSVTVSITSGASENSCTVEKKIR
 RKFVGGREEYLFPPVHGKLVKCHVYDHLKETSAGYITMHRPGPHAYKSYLEEASGEVYI
 KPPSGKNVITYECKG DYSTGIVSTRTKMNGCTKAKQCIAYKSDQTKWVFNSPDLIRHT
 DHSVQGLKHIHPFRLTPTVCVPPLAHTPTVTWKFGITLHLTAMRPTLLTTRKLGRLAD
 ATAEWITGSTSRNFSVGREGLYVWGNHEPVRVWAQESAPGDPHGWPH EIIHYHHRH
 PVYTVIVLCGVALAILVGTASSAACIAKARRDCLTPYALAPNATVPTALAVLCCIRPT
 NAETFGETLNHLWFNPNQPFILWAQLCIPLAALVILFRCFSCCMFLLVAGVCLGKVDAF
 EHATTVPNVPGIPYKALVERAGYAPLNLEITVVSSELTPTSTNKEYVTCKFHTVIPSQ

1003619.122101

VKCCGSLECKASSKADYTCRVFGGVYPFMWGGAQCFCDSENTQLSEAYVEFAPDCTID
 HAVALKVHTAALKVGLRIVYGNTTAHLDTFVNGVTPGSSRDLKVIAGPISAAFSFPDH
 KVVIRKGLVYNYDFPEYGAMKPGAFGDIQASSLDTDIVARTDIRLLKPSVKNIHVPI
 TQAVSGYEMWKNNSGRPLQETAPFGCKIEVEPLRASNCAYGHIPISIDIPDAAFVRSS
 ESPTILEVSTVADCIYSADFSSSLTLQYKADREGHCPVHSHSTAVLKEATHTHTAV
 GSITLHFSTSSPQANFIVSLCGKKSTCNAECKPPADHIIGEPHKVDQEFQAAVSKTSW
 NWLLALFGGASSLIVGLIVLVCSMLINTRR"

3' UTR

3870-4150

BASE	COUNT	ORIGIN
1	ggccctctag	agctcatact
61	agcggttaaga	acttcaagag
121	taggtgacgt	agtagacacg
181	actttccacc	agtttaccct
241	gccgctggag	gccgtttcgg
301	tagtcaactt	gactttcaaa
361	agaagaagag	tgctoctaag
421	agaggacgaa	acgcaagcct
481	cggacaagac	atttccgac
541	gaggaaggct	gatgaaacca
601	ccgtgaaatt	gaagaaggct
661	tgaatcaga	cacgctgcag
721	acggcgacgt	ccagtatgag
781	cgcacacgcg	aagaccgatc
841	gtgcaaatga	gggcacgcgt
901	ccattaggga	tacccccgaa
961	tttcgaatgt	cacgtttccca
1021	gaacactcga	cgtgtcgaa
1081	acgtcttgaa	gtgtccatca
1141	ccagtcacct	cctgggggtc
1201	taaaaaattga	gaagcttggg
1261	caacaattcgg	ctacaatcag
1321	tcgaccacga	ccatgacatc
1381	gacctctgcg	tcgtcttggc
1441	acagtgtaac	cgtcagtatc
1501	agatcaggag	gaagtttgct
1561	tggtaaagtg	ccacgtttac
1621	acaggccagg	cccacacgcg
1681	ttaaacacc	ttctgcgaag
1741	gtatcgtgag	cacgcgaacg
1801	acaagagcga	ccaaacgaaa
1861	actcagtgca	aggtaaattg
1921	cgtttagctca	cacgcctaca
		gtcacgaagt
		ggcaggcctg
		cataagaggg
		cacctaccca
		acaaatccga
		ccccgcgtgg
		caacgatcac
		ccaaaaccta
		aaaccaggga
		atgctgaacg
		ctccacgttg
		agcatgtacg
		tacaccagcg
		aatgggagat
		ctgggacaaca
		acggcgcttt
		ggttctgaac
		tgcgacaaac
		gagaacgtcg
		cgcgggcccc
		tgcccgctatt
		gcagacactc
		atgatgctcg
		gcaggcactg
		aaggaagaca
		acaaaaggg
		acgagcggag
		ggtagagagg
		gatcacttga
		tataagtctc
		aacgtcacct
		aaagatgaacg
		tggtgtcttca
		cacattccat
		tccgcttgac
		ggttcaaagg
		atcatcacgt
		agcccaatca
		ccctctacgg
		ccggcagaat
		tggtctaccg
		ctgctcaaat
		ctaatcgcc
		cccgaggtcca
		aaagaagaag
		atgtgtgatg
		tgccgtgtcg
		tgaatagtag
		cggcgacgtt
		gggcttctac
		gagaggagtg
		tggtgctatt
		tgggaaccag
		agttacagcg
		ctattcactg
		taccagacg
		taccgatgac
		aacgcgtgtg
		gattagaatc
		caaatccctg
		aatagctatc
		acttctctgt
		ttcatgcacc
		cccacccgtc
		tgccgggtac
		agcgtcaggc
		gtgtggcgac
		agcaaaaacg
		ctctattagg
		accgacagtc
		catcaccttc
		gttagccgaa
		ccctcagctga
		cctcctagggc
		aggaggtcga
		ccgccaaaaga
		cagcaagcca
		aagttggagt
		tgccgtgtcg
		caattagcgg
		ccccagaaca
		aactggccacc
		ggcgggaaag
		gttctaggag
		aaaggggtga
		ctatgcgtgc
		acgcagaaac
		ctgctggaga
		ttcacactga
		ttcagcccaa
		caggtctcgg
		tacatgtctt
		agcacatctg
		gtggagaaaa
		caggtctcgg
		ataaccatgc
		gaagtgatca
		tacagcacag
		tgcatgtcct
		cacacagacc
		tgcccggttc
		cacctgactg

```

1981 caatgcgacc aacattgctg acaacgagaa aattggggct gcgagcagac gcaacagcag
2041 aatggattac aggggtctaca tccaggaatt tttctgtggg gcgagaaggg ctggagtagc
2101 tatggggtaa ccatgaacca gtcagagtct gggcccagga gtcggcacca ggcgacccac
2161 atggatggcc gcatgagatc atcatccact attatcatcg gcatccagtc tacactgtca
2221 ttgtgtctgt tgggtgcgct cttgtctatcc tggtaggcac tgcatcatca gcagcttgca
2281 tcgccaaggc aagaagagac tgcttgacgc catacgcgct tgcaaccgaac gcaacgggtac
2341 ccacgacatt agcggttttg tgctgcattc ggccaaccaa cgctgaaaaca tttggagaaa
2401 cttttgaacca tctgtggttt aacaaccaac cgtttctctg gcacagttg tgcatctctc
2461 tggcagcgct tgttattctg ttccgctgct tttcatgctg catgcctttt ttattgggtg
2521 caggcgctgt cctggggaag gtagacgctt tcgaacatgc gaccactgtg ccaaatgttc
2581 cggggatccc gtataaggcg ttggtcgaac gcgcaggtta cgcccaactt aaactggaga
2641 tcacggctcg tcatcgggaa ttaacacctt caactaacia ggagtacgtg acctgcaaat
2701 tccacacagt cattccttca ccacaagtta aatgctgcgg gtccctcgag tgcaaggcat
2761 cctcaaggcg ggattacaca tgccgcgctt ttggcgggtg gtaccttttc atgtggggag
2821 gcgcacaatg cttctgtgac agtgagaaca cacaactgag tgaggcgtag gtcgagttcg
2881 ctccagactg cactatagat cacgcagtcg cactaaaagt tcacacagct gctctgaaag
2941 tcggcctcgg tatagtatac ggcaacacca ccgcgcacct ggatacgttt gtcfaatggcg
3001 tcacgccagg ttctctcagg gacctgaagg tcatagcagg gccgatatca gcgccttttt
3061 caccctttga ccataaagtc gtcatacagaa aggggcttgt ttacaactac gacttccctg
3121 agtatggagc cttatggaca ggagcgctcg gcgatattca atcatcctcg cttgatgcta
3181 cagacatagt agcccgact gacatacggc tgctgaagcc tctgtcaagg aacatccacg
3241 tccctacac ccaagcagta tcagggtatg aaatgtggaa gaacaactca ggacgacccc
3301 tgcaagaaac agcaccattt ggaatgtaaaa ttgaagtgga gctctgcgga cgttctaact
3361 gtgcttacgg gcacatccct atctcgattg acatccctga tgcagctttt gtgagatcat
3421 cagaatcacc aacaatttta gaagttagct gcacagtatg agactgcatt tattctgcag
3481 actttgggtg ttctctaaac ttacagtaca aagctgacag ggagggacat tgtccagttc
3541 actcccactc cagcagagct gttttgaagg aagcgaccac acatgtgact gccgtaggca
3601 gcataaacat acattttagc acatcgagcc cacaagcaaa tttttatagt tcgctatcgg
3661 gcaagaagtc cacttgcaat gctgaatgta aaccacggcg cgaccacata attgaggaaac
3721 cacataaagt cgaccaagaa ttccaggcgg cagtttccaa aacatcttgg aactggctgc
3781 ttgcactggt tggggggagca tcatccctca ctgtgttcag actttatagt ttggctcgca
3841 gctctatgct tataaacaca cgtagatgac tgagcgcgga cactgacata gcggtaaaaa
3901 tcgatgtact tccgaggaag cgtgggtgcat aatgccacgc gccgcttgac actaaaaactc
3961 gatgtatttc cgagggaagca cagtgcataa tgctgtgcag tgtcacatta atcgtatatc
4021 aactacataa ttaacaacac tatatcactt ttatgagact cactatgggt ctctaataa
4081 cactacacat attttactta aaaaacatat acacacttta taaattctct cataatttoa
4141 ctttaggttt

```

//

SEQUENCE LISTING - pVHX-6 - SEQ ID NO 3

<200> SEQUENCE CHARACTERISTICS
 <210> SEQ ID NO 3
 <211> LENGTH: 1260
 <212> TYPE: DNA
 <213> ORGANISM: EXAMPLE OF A CMV PROMOTER COUPLED TO THE pcDW-XH7
 SEQUENCE
 <220> FEATURE:
 1-302 pVAX vector sequence
 302-401 CMV promoter
 411 CMV putative transcriptional start site
 444-463 T7 promoter
 464-509 pVAX multicloning region
 510 start of pcDW-HX45 sequence

1	cattgacgtc	aatgggtgga	gtattttacgg	taaactgccc	acttggcagt	acatcaagtg
61	tatcatatgc	caagtacgcc	cctatttgac	gtcaatgacg	gtaaatggcc	cgcttggcat
121	tatgccagtg	acatgacctt	atgggacttt	cctacttgcc	agtacatcta	cgtattagtc
181	atcgctatta	ccatggtgat	gcggttttgg	cagtacatca	atgggcgttg	atagcggttt
241	gactcagcgg	gatttccaag	tctccacccc	attgacgtca	atgggagttt	gttttggcac
301	caaaatcaac	gggaactttc	aaaatgtcgt	aaacaactccg	cccatttgac	gcaaatgggc
361	ggtaggcgtg	tacgggtggga	ggcttatata	agcagagctc	tctggctaac	tagagaaccc
421	actgcttact	ggcttatcga	aattaatcag	actcactata	gggagaccca	agctggctag
481	cgtttaaaact	taagcttggg	accgagctca	tactggcagg	cctgatcatc	acgtctctgt
541	ccacggttagc	cgaagcgtt	aagaacttca	agagcataag	agggagccca	atcacctct
601	acggctgacc	taaatagggtg	acgtagtaga	cacgcacctta	cccaccgccca	gaatgtttcc
661	ataccctcag	ctgaactttc	caccagttta	ccctacaaaat	ccgatggctt	accgagatcc
721	aaaccctcct	aggcgcgcgt	ggaggccggt	tcggccccc	ctggctgctc	aaatcgaaga
781	tcttaggagg	tcgatagtca	acttgacttt	caaacacaga	tcacctaatc	cgccgccagg
841	tccaccgccca	aagaagaaga	agagtgtctc	taaggcaaaa	cctactcagc	ctaaaaagaa
901	gaagcagcaa	gccaaagga	cgaaacgcga	gcctaaacca	gggaaacgac	aacgtatgtg
961	tatgaagtgtg	gagtcgggaa	agacatttcc	gatcatgtcg	aacggccaag	tgaatggata
1021	tgcctgcgtt	gtcgggagaa	ggctgatgaa	accactccac	gttgaaggaa	aaattgataa
1081	tgagcaatta	gcggccgtga	aattgaagaa	ggctagcatg	tacgacttgg	agtacggcga
1141	cgttccccag	aaactgaagt	cagacacgct	cgagtagacc	agcgacaaa	caccggggct
1201	ctacaaactg	caccacggcg	cagtcacgta	tgagaatggg	agatttaccg	taccgagagg

. . . . CONTINUOUS WITH pcDW-HX45 (pcDW-XH7) SEQUENCE

4177	gaatgtaaac	cacogggcca	ccacataaatt	ggagaaccac	ataaaagtcca	ccaagaattc
4237	caggcggcag	tttccaaaac	atcttggaac	tggtgtgctt	cactgtttgg	gggagcatca
4297	tcctctcatt	tttgtaggact	tatatgtgtg	gtctgcagct	ctatgcttat	aaacacacgt
4357	agatgactga	gcgcggacac	tgacatagcg	gtaaaactcg	atgtaacttc	gaggaagcgt
4417	gggtgcataat	gccacgcgaa	gcttgacact	aaaactcgat	gtatttccga	ggaagcacag
4477	tgcataatgc	ttgtgcagtgt	cacattaatc	gtatatcaca	ctacatatta	acaacactat
4537	atacactttta	tgagactcac	tatgggtctc	taatatacac	tacacatatt	ttacttaaaa
4597	acactataca	cactttataa	attcttttat	aatttttctt	ttgctttaga	gcacactggc

4657 ggccgttact agtggatccg agctctagag ggcc

4177-4643 pcDW-HX45

4644-4790 vector sequence //

101221-412300